International Conference on Artificial Reality and Telexistence
Eurographics Symposium on Virtual Environments (2016)
D. Reiners, D. Iwai, and F. Steinicke (Editors)

Is this seat taken? Behavioural analysis of the Telethrone: a novel situated tele-presence display.

J. O'Hare[1] and R. C. A. Bendall[2] J. Rae[3] G. Thomas[4] B. Weir[4] D. J. Roberts[2]

[1]University Of Salford, Computer Science & Engineering, U.K.
[2]University Of Salford, School of Health Sciences, U.K.
[3]Roehampton University, Department of Psychology.
[4]BBC R&D, Media City, Salford, UK.

Abstract

We present research with two novel components; a system which may improve current small group telecommunication, and an experiment to test the efficacy. Telethrone projects a remote user onto a chair, bringing them into your space. The chair acts as a situated display which can support multi party head gaze, eye gaze, and body torque such that each observer knows where the projected user is looking. It is simpler to implement and cheaper than current systems. Our primary contribution is a counterbalanced repeated measures experiment to analyse gaze interactions. We analyse the multiple independent viewpoint support offered by the system to test if it demonstrates advantage over a set-up which shows a single view to both observers; in this results are inconclusive. Self-report questionnaire data suggests that the current implementation still gives the impression of being a display despite its situated nature although participants did feel the remote user was in the space with them. Results from the eye gaze analysis suggest that the remote user is not excluded from three way poker game-play.

1. Introduction

Wouldn't it be nice to have natural conversations with someone in another office, home, or country, without feeling that technology was somehow 'in the way'? Video conferencing (VC; 'Skype', in its best known form) has made real inroads in supporting this, but looking through a screen keeps the other party very much *in their space* as in Triple-View which employs computer monitors showing windows into another space [Can11]. Unlike VC, situated displays [PS14] attempt to put the remote person *inside your space*, while ensuring that gaze and gesture can retain spatial context.

We describe the Telethrone: a novel situated display that places a representation of a remote participant in an actual chair within the user's space. This maintains a natural context, in contrast to other contemporary research systems. We suggest that visualizing the whole of the remote user may be less weird than an anthropomorphic tele-robot or 'head in a jar' style approach. The initial Telethrone design was a response to a problem BBC Research and Development (R&D) encountered through split site working whereby they could not 'rub shoulders' over coffee. In addition, BBC R&D suggested that 10 unstructured *ad hoc* meetings were optimal for leveraging the best outcomes from a single formal meeting. In a single site environment, such as their original London headquarters, this kind of meeting was supported through communal areas and the eponymous 'water cooler meetings'. With the advent of split site working in London and Salford it became apparent that these clarifying meetings were being compromised. The ideal solution therefore would be a technology which could be economically integrated into a social space and could be 'always on'. Serendipitous meetings through a screen on a wall had been attempted by BBC R&D but did not seem to work, even if the system were always on. One limitation may be the difficulty in correctly grabbing the attention of passing colleagues due to the spatial misalignment inherent in the Mona Lisa effect [AMEB12] and the limitation of the flat screens 'containing' another space.

The Telethrone can provide independent spatially correct views from positioned seats, supporting multi-directional channels of gaze and body torque, and other non-verbal communication (Figure 1 & 2).

We investigate the degree to which this comparatively simple system represents the remote user. We seek to bring an impression of the whole of the remote user into the space in a natural way, without noticeably drawing in aspects of their environment with them. We believe this support for affordable, natural setting multi-view is poorly supported by current available commercial and research systems.

2. Related Work

Point to point connection of multiple users, with IP network encoding of voice and or facial camera, is well supported through

Figure 1: *View from the participant toward the Telethrone*

Figure 2: *View from the experimenter toward the Telethrone*

technologies such as telephony & Skype, with Skype alone accounting for 280 million connections per month [Mun13]. However, these technologies demonstrate shortfalls compared to a live face-to-face meeting, which is generally agreed to be optimal for human-human interaction [WRM*08] [Var02]. Support for mutual eye gaze, and spatially faithful transmission of body torque in multi-party sessions are two important gaps in these technologies [BBBL01]. Face-to-face communication can therefore be considered the benchmark by which the approaches detailed below should be judged. These established technologies extend traditional telephony to provide important multi-model (multiple sense) cues through non-verbal communication [AG76] [WRM*08].

2.1. Video Conferencing, Gaze, and Body Torque

Single user-to-user systems based around bidirectional video implicitly align the user's gaze by constraining the camera to roughly the same location as the display. When viewed away from this ideal axis, it creates the feeling of being looked at regardless of where this observer is (the Mona Lisa effect). Multiple individuals using one such channel can feel as if they are being looked at simultaneously, leading to a breakdown in the normal non-verbal communication which mediates turn passing [VD02].

Some non-verbal communication is supported in VC with limited success. Additional screens and cameras add multi-party support to a degree. This mitigates the problem of addressing a room full of people from a single screen by making available more bidirectional channels. Every remote user can be a head on a screen with a corresponding camera. The positioning of the screens must then necessarily match the physical organization of the remote room or rooms. Supporting spatial aspects such as mutual gaze in this way therefore demands large purpose built installations which poorly support casual or ad hoc meeting paradigms [SSA*01] [WRM*08]. Nonetheless these systems represent a healthy and growing industry.

However, most of these conventional single, and expensive multi camera VC systems, suffer a fundamental limitation in that the offset between the camera sight lines and the lines of actual sight introduce incongruities that the brain must compensate for [WRM*08]. Gaze-2 [VW03] overcame this problem using simple video channels by tracking eye movement and switching camera feeds from multiple cameras to all the remote connections. This ensured correct 'one to many' relationships with the remote users. The Multiview system [Ngu05] also demonstrates spatial faithfulness and demonstrated increased trust and persuasion between connected groups compared to traditional VC [NC07].

While it is well proven that there are advantages to accurate connection of the gaze between conversational partners [AI69] [Kle86], there is also a body of evidence that physical communication channels extend beyond the face [Kle86] [NC09] and include both micro (shrugs, hands and arms), and macro (torque) movement of the upper body [Ekm93].

2.2. Interpolated View and View Reconstruction

Triple-view [Can11] affords correct spatial alignment through the use of an interpolated camera pair per screen, providing three bidirectional channels which preserve the directionality of the user's gaze. Importantly this system still uses screens which show backgrounds from the remote space resulting in the feeling of looking into another space. The more technically demanding Immersive Group-to-Group [BKKF13] wall places the reconstructed remote users in a bland virtual space which can be set to match the surrounding walls and draws less attention to the elements which are not the remote collaborator.

2.3. Situated Displays

Between the complexity of ICVE's and the more ubiquitous screen based VC technology there now exist situated displays. These place

a representation of the remote user into a space and are a relatively new field of research. Embodiment through hybridisation of real-time video and physical animatronic mannequins has been investigated as a way to bring the remote person into the space in a more convincing way [LWN*09] [RWLB01] [STE*05]. These include Tele-presence robots [LT11] [SKO*07] [TDYU11], head in a jar implementations such as SphereAvatar [OSS12] [PSS14] [PS12], the 'Gaze Preserving Situated Multi-View Telepresence System' [PS14], or screen on a stick style representations [KCL13]. Telehuman brings the whole body of a standing remote user into a space via a cylindrical display with a single tracked observer viewpoint [KBG*12].

2.4. Summary

It is clear that there is a justification for a step change in simple affordable technologies which better mediate communication over distance (as expounded in Distributed Work [Var02]). Reconnection of naturalistic non-verbal cues bolsters turn passing, trust, empathy, and rapport between co-located, and tele-present users [BHPR]. The Telethrone aims to address this requirement for a simple, deployable, pervasive, group telecommunication system with spatial, non-verbal cue support. Figure 3 summarises some of the important referenced systems including the Telethrone and compares their features or 'affordances'.

Systems	Joint eye gaze	Situated	Faithful spatial body torque	3D	Reconstruct viewpoint	TCVE	Multi view points	Natural setting
Skype	No	No	No	No	No	No	No	No
Multiview	Yes	No	Yes	No	No	No	Yes	No
Majic	Part	No	Yes	No	No	No	No	No
GAZE-2	Yes	No	No	No	No	No	No	No
Sphere avatar	Yes	Yes	No	No	Yes	No	Yes	No
Withyou	Yes	No	Yes	Yes	Yes	Yes	No	No
Tele-human	Yes	Yes	Yes	No	Yes	Yes	No	Part
Immersive group to group	Yes	Part	Yes	Yes	Yes	No	Yes	Yes
Triple-View	Yes	No	Yes	Yes	Yes	No	Yes	Part
Cisco TP	Part	No	No	No	No	No	No	Yes
UCL situated	Yes	Yes	No	No	No	No	Yes	No
Telethrone	Yes	Yes	Yes	No	No	No	Yes	Yes
Room2Room	Yes	Yes	Yes	No	Yes	No	No	Yes

Figure 3: *important referenced systems including the Telethrone*

3. Technology Development - Telethrone

We introduce the Telethrone, a projection system which attempts to situate a remote user on a multi-view chair. The solution uses low cost commodity components to address the requirement for an 'always on' technology which can be deployed in an everyday setting. The Telethrone system is 'situated', which may be a more comfortable and natural mediator of remote telecommunication. Many attempts to encourage ad-hoc meetings through video conferencing have largely failed and no solution has yet gained wide acceptance. Even in an arranged meeting, there is something social about sitting next to a person in a chair. A flat screen, even if it could overcome the Mona Lisa effect, would at best be like meeting someone through a sheet of glass. Unlike a framed screen, the Telethrone occupies the space at human scale, in the familiar context of a chair. The system attempts to isolate the remote user, cutting most of the background from the projection and filling the chair. This balances the meeting physically with life-size human forms, and retains the desirable spatial faithfulness and multi-view demonstrated successfully in other systems [Ngu05] [NC07] [Zha07].

3.1. Multi-view through Chromatte

Multi-view Telethrone supports gaze direction and body torque [Can11] by providing independent viewpoints to multiple users. Reflecmedia Chromatte is a commercially available cloth with retro-reflective property, that is, it tends to reflect light back in a cone along the angle of incidence as seen in Figure 4. It is essentially the same material high visibility jackets are made of. The Telethrone is a high backed chair draped in this cloth. The retro-reflective property of Chromatte is useful in spatially isolating projections. This usage has been termed "retro reflective projection technology" or RPT [Tac03]. The Tele-throne employs this to spatially separate video streams. Two video feeds from the cameras in a remote room were projected and retro-reflected from the Chromatte cloth, bouncing back to the two local participants as in Figure 1 and Figure 2.

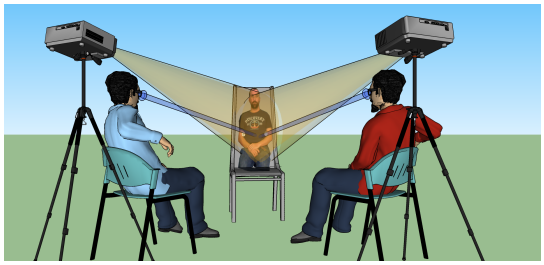

Figure 4: *Light from the projectors bouncing back along the angle of incidence.*

There is a cross-talk effect between the two spatially separated images which minimizes after around 15 degrees (commensurate with manufacturer's datasheets), but which otherwise remains at a value of around 5%. Potential effects of this characteristic were thereby a consideration throughout the investigation, i.e., would a slight double image seem weird, or more specifically be testable as uncanny [Mor70, GBK05, MSSL*11].

Figure 5: *View between the two observers where cross-talk between the spatially segmented views is apparent.*

4. Experiment

4.1. Set-up

Video was swapped between two rooms to form a Tele-presence connection. Two cameras (on tripods) by the large TV on the left

of Figure 2 were transmitted directly to two projectors mounted behind the locally situated users on the right of Figure 6. The retro-reflection from the projector frustum to the eyes of the observer is illustrated in Figure 4

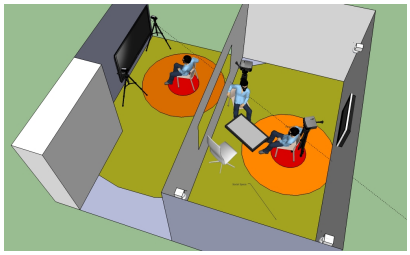

Figure 6: *Two locations can be seen on the left and right of the imageThe Telethrone is shown as an empty chair.*

The figure above illustrates two adjoining rooms connected by a one way mirror which hides the cameras to reducine the observer effect [Arg] [HH69]. The diagram also illustrates the proxemics social space in green, with personal space in orange and intimate in red.

4.1.1. Observation Room & Camera Connections

The remote room to the left of Figure 6 consistently hosted a tele-present researcher. The researcher sat facing an LG 60PF95 plasma TV whose image showed the other researcher and the participant seated in the observation room (Figure 7).

Figure 7: *The remote user's station with IP cameras*

We attempted to maximize spatial accuracy, replicating the triad in the observation room for the remote observer. This transmitted the correct angle of head rotation from the remote person when their attention switched from participant to researcher or vice versa. The offset from the center of the eyes on the TV screen to the camera was roughly 5 degrees, leading to a 5 degree deviation from true on the returned gaze.

The low latency (˜<10ms) direct connections provided around 20fps constrained by the brightness of the lighting in the room which was in turn limited by the relative brightness of the available projectors. Faster projectors would potentially enable a higher frame-rate up to the 48fps maximum of the cameras. We did not seek to emulate the delay of a typical tele-presence set-up, as this would present a further confounding factor.

The projectors were height adjusted per experimental run (on tripods) such that their projection frustums sat just above the heads of the participant and researcher who shared the observation room. This tuning per session accommodated height differences between participants and can be seen demonstrated with the base of a projector just above the head of the researcher as in Figure 2.

4.1.2. Single-view

The experiment compared gaze behaviour of the participant between a co-located researcher and a remotely located projected researcher. Two types of remote projection technique were tested. Of primary interest was multi-view, theoretically supporting directionality and mutual gaze between participant and tele-present researcher. Also investigated was single-view which employed an offset camera (centre of the TV) such that the view of the tele-present researcher would be impossible to reliably resolve (having more than 10 degrees horizontal offset).

In our single-view configuration the video from a camera in the centre of the TV was linked to both projectors simultaneously, and then these video images were aligned to one another on the Chromatte. In this mode both local users see the transmission from this same camera, situated directly in front of the remote person. This created a false off-axis view similar to the spatial offset problem demonstrated by commodity VC systems.

It was predicted that natural communication would be maintained with spatial (directional) non-verbal cues being supported despite the use of the technology. It was expected that differences in gaze behaviour would be evident between Telethrone single-view and both the physically co-located and the Telethrone multi-view conditions. We predicted that fewer differences in gaze behaviour would be evident between Telethrone multi-view and co-located conditions.

4.1.3. Audio

Audio was transmitted to the remote person in stereo from two directional ClearOne desk condenser microphones. These were stationed under the central table close to the users in the room, and sent to discreet headphones visible in Figure 7. This stereo signal was given phantom voltage and routing using a Yamaha 01v mixing desk which also served to power a single condenser microphone. This microphone in the remote room delivered the voice of the tele-present researcher from the remote to a Behringer CE500A powered monitor speaker positioned directly behind the projected head(s) on the Telethrone, creating the impression of their voice coming from the head area of the projection.

4.2. Methoddology

4.2.1. Hypotheses

- H1: That the remote participant is not excluded from the conversation through significantly less attention as measured by examination of eye gaze events.
- H2: That the multi-view condition demonstrates more natural looking behaviour than the single-view condition as examined through eye gaze events and questionnaire .

4.2.2. Participants

Sixteen participants (14 male) aged between 18 and 46 years (M = 31.89 years, SD = 8.5 years) participated in this study. Written informed consent was gained from each participant after they were given procedural information about the study. Ethical approval was obtained from the School of Computer Science and Engineering Research Ethics Panel at the University of Salford (CST 15/03). All participants received a Âč20 inconvenience allowance.

4.2.3. Design

A within-participants design was used with two independent variables; the medium of communication (co-located or projected onto the Telethrone), and the support for directional view. The dependent variables were the number of look events (glances and looks), and total duration in seconds of participant gaze to both the co-located and the Telethrone remote researcher. Responses were recorded from repeated self-report questionnaires for both the multi-view and the single view conditions.

4.2.4. Materials

Participants completed tele-presence questions (e.g. ''I felt that the person on my left/right was in the space with me?'') between projection conditions and at the end of the sessions. Participants answered the questions on a 7 point Likert scale ranging from 1 (not at all) to 7 (very much). To attempt to assess potential uncanny effects of the double image from the projectors we used the likeability and anthropomorphism sections from the GODSPEED uncanny valley questionnaire [6] administered using a 7 point Likert scale ranging from 1 (machinelike) to 7 (humanlike).

4.2.5. Procedure

Participants started the study after reading an information sheet and providing informed consent. The experiment separated three people by putting two together in a common room, with the third remote person in an adjacent room. The remote researcher interacted with a TV showing the two people in the common room; while those in the common room viewed the remote person through the medium of the Telethrone. Two of the users were confederate researchers. One researcher was located in the remote room whilst the other research was located in the observation room. The other user was a participant.

The experimental set-up reflects the 2m optimum distance identified by Hall [HH69]. The participant is seen on the right of 6 in the proxemics rings with the other players in the 'social space' banding. The three seats in the main room were distributed evenly around the table to balance the conversation spatially.

Shared common tasks are a prevalent feature of group tele-presence research as this more accurately reflects professional or social group meetings. A structured task also enhances repeatability under experimental conditions. A disadvantage of a less familiar task is that it can radically reduce eye contact and especially mutual eye contact. More structured group problem solving tasks promote turn taking and thereby support gaze however they are complicated to create and analyse [VVV00]. A card game was implemented in computer graphics to simplify the design. The card game was designed for a touch table (Figure 8). Poker was chosen as the shared task as it is a familiar group activity to many. Additionally, it may be that poker bolsters observation of other players (reading a players bluff). It was established by ÃŸstrem [Ost14] that the act of gambling for money increases emotional engagement, so we ensured a financial incentive in our study design. The card game was implemented on touch table in Flash/ActionScript.

The experimental design closely matched the physical layout of a real game through the use of the 42 inch touchscreen display mounted horizontally at the same height as a normal card table. Player's cards were dealt face up under a physical mask that closely matched the positioning of cards which would be face down in a real game (Figure 2). To look at the cards the players were forced to physically move. This movement resembles the action employed in physical play, where a player will lift the edges of their own cards and duck slightly to check their values.

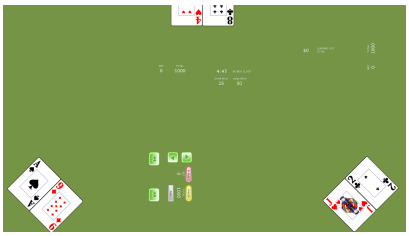

Figure 8: *A screenshot from the poker game*

Buttons available on the touchscreen display in communal view allowed bets to be increased and decreased. They also facilitated folding and the passing of the locus of control. Player chip values and community pots were shown in communal view as in a traditional game.

The distributed poker game allowed game infractions in the same way a physical card game might (i.e. cards could be dealt at the wrong time, players could make out of turn actions). Claiming chips was also a manual process rather than automated. These incidents and interactions promoted discussion during the game through observation of specific actions and additional attention to other players' behaviour.

The tele-present player interacted using a laptop on their knees with the screen pushed back to near horizontal. This closely spatially approximated the 42" touch table. Physical masks taped to the laptop screen covered the other players' hands. The set-up was faithful to a genuine poker game, with no cheats or biases for the experimenter.

Two five-minute practice games were completed which combined instruction and practice during which conversation was encouraged (typically initiated by the researchers). After this participants completed two rounds of ten minutes on one technology (multi-view or single-view). They then completed the questionnaire comparing co-located researcher with the Telethrone projection. The projection set-up was swapped during this period, and the next two rounds were played before an additional - and final - repeat of the questions.

Chip totals were summed on paper between hands with eventual 'chip leader' initially agreed to win Âč20 and the other two players Âč5. In actuality each subject was thanked and awarded the full Âč20 at this stage.

5. Analysis and Results

To prepare for analysis, the 6 video streams were synchronized. Sections of game play (i.e. not briefing or interruptions for scoring or breaks) were isolated, presenting a condensed dataset for inspection.

Analysis was initially based upon continuous visual inspection of the video streams (predominantly from a single 'over the shoulder' view use employed in Hall [AI69]). Glances from the participant toward what appeared to be faces for either co-located or telepresent players were counted. This method was repeated for consistency by another researcher. Next a deeper *frame by frame* analysis was undertaken marking from start of glance to termination of glance for all glances. Anything deemed to be directed toward one or the other players faces was marked in Cinelerra-CV for the duration of the look event. Where the glance was ambiguous inspection of multiple camera angles was undertaken. Removing the gaps between these marks gave a total time for 'looking' for each session of play.

The relatively small sample size of n = 16 meant that we could not test if the data were parametric, so we chose to apply Wilcoxon signed rank tests to the data.

5.1. Eye gaze results

For both multi-view and single-view projection conditions there were more looks (gaze events) toward the Telethrone than toward the co-located researcher.

In the multi-view projection condition the number of gaze events toward the Telethrone were significantly greater (*Mdn* = 33) than the co-located researcher (*Mdn* = 22.35), T = 22, p = .017. In the single-view projection condition there were more gaze events toward the Telethrone (*Mdn* = 31) than the co-located researcher (Mdn = 22.35), T = 113, p = .020. In comparing projection conditions there is no significant difference between gaze events toward multi-view vs gaze events toward single-view, T = 57.5, p = .587. There is no significant difference between single-view and multi-view for number of look events when compared to one another, T = 78.5, p = .587.

In both projection conditions there was increased gaze duration toward the Telethrone than the co-located researcher. In the multi-view projection condition the total duration of gaze toward the Telethrone was significantly higher (*Mdn* = 29s) than for the averaged co-located researcher (*Mdn* = 23.28), T = 23,p = .020, while for the single-view projection condition gaze duration toward the Telethrone was significantly higher (*Mdn* = 39.88) than for the averaged co-located researcher (*Mdn* = 23.28), T = 92.5, p = .012. There is no significant difference between single-view and multi-view for look duration when compared to one another, T = 67, p = .959.

5.2. Questionnaire responses

	co-located researcher	Telepresent researcher					
		Single-view			Multi-view		
Question	*Mdn*	*Mdn*	T	*p*	*Mdn*	T	*p*
I felt that person was in the space with me	7	5	102	.002	6	105	.001
I felt I was looking at the person through a screen	1	4	2.5	.004	3	115	.003
It seemed to me the person occupied a different space to me	1	3.5	1	.002	3	136	.001
when I looked at the person it seemed I was looking into another room	1	2	2.5	.012	2.5	77	.003

Figure 9: *Results from the questions*

Figure 9 shows responses to questions which attempt to explore how situated the display is and how much the remote collaborator becomes part of the space. They are statistically significant responses.

Additionally there are interesting non significant differences between the physically co-located researcher and the tele-present researcher and across the projection conditions.

In response to the question "I felt I was in the same room as two other people" there was no significant difference between multi-view vs single-view projection modes, T = 33.5, p = .964.

In response to the question "During the game I felt I was playing with two other people" there was no significant difference between multi-view (*Mdn* = 7) vs single-view (*Mdn* = 7) projection modes, T = 7, p = .891.

In response to the question "I felt that the person on my left interacted naturally with the person on my right" there was no significant difference between multi-view (*Mdn* = 5) vs single-view (*Mdn* = 6) projection modes, T = 44.5, p = .943.

The GODSPEED anthropomorphism section found no significant differences between multi-view (*Mdn* = 5) and single-view (*Mdn* = 6) projections conditions T = 410, p = .309. The GODSPEED likeability section found no significant differences between multi-view (Mdn = 6) and single-view (*Mdn* = 6), T = 460, p = .696. Overall across all GODSPEED indicators (*Mdn* = 6).

6. Discussion

The aim of the current study was to investigate if the remote participant was excluded from the conversation as measured by looking / eye gaze events. We demonstrate that not only were tele-presence participants included but that the Telethrone attracted more attention than the co-located participant. This was true for number of gaze events in both projection modes, as well as duration of gazing in both projection modes.

The reasons for this general bias toward the Telethrone remain unclear. We suggest that the physical flow of play may give rise to increased glances at the Telethrone as play is always to the left, and the Telethrone always to the left of the participant. It may also be the novelty of the Telethrone. It may be inherent weirdness, or the compounding weirdness of the slight double image stimulates

additional looking. It is also conceivably a function of the subjects taking more time attempting to resolve the attention of the remote player. This is potentially suggested in the data in that the deliberately poor single-view is looked at the most (*Mdn* = 39.88) with multi-view second (*Mdn* = 29s), and the averaged co-located researcher the least (*Mdn* = 23.28). This difference between multi-view and single-view conditions is not significant.

We also investigated whether the Telethrone supported directionality of gaze, mutual gaze, and body torque. The difference between these conditions was unclear. Had such a difference been evident in the data it may have been possible to assert that the system also supported mutual gaze in the multi-view system, and thereby would have supported it as a system above and beyond conventional approaches. That our investigation did not expose such a difference is somewhat at odds with findings from the Triple-View system, which compared their spatially faithful system with a similarly cut-down version to the single-view mode of Telethrone. Triple-view suggested that gaze direction was important when they employed a collaborative task and analysis of turn taking while solving language puzzles [VVV00]. Vertgaal et al also found that there was a slight increase in all of their reported questionnaire metrics for mutual gaze [Can11]. It may then be that the contrived and competitive nature of poker, especially the 'covert' aspect of some of the observation gives rise to unnatural look patterns. It was observable in our analysis that there was very little mutual gaze, with competitors preferring observational glancing to judge game-play over conversational and/or communicational glancing and looking.

Median overall modified GODSPEED indicators were 6 on a 1-7 scale where 1 demonstrates a representation of a person most 'uncanny' and 7 most human. This suggests that the system was not particularly uncanny according to the tested criteria, with no significant difference between projection conditions. This is somewhat interesting in that small affects to the representation of a human through technology can have disproportionate effects on the impression the representation gives [Mor70], and this was not observed.

Recently Microsoft labs announced that they had been researching projection onto furniture using their Kinect system and projectors [TP16]. This lends weight to our opinion that projection onto furniture is a potentially exciting and important area of research. Their system provides spatially correct viewing through reconstruction but appears only to work on a point to point basis with a single user at each end. Telethrone offers significant advantage through its support for multiple viewpoints.

7. Conclusion

We have presented the Telethrone which we suggest offers an advantage over current research situated displays through combination of affordability, technical simplicity, contextual fit and natural look. Our analysis supports our central hypothesis that the display does not exclude the remote participant from the conversation, but currently fails to find significant advantage to multi-view as compared to the single-view projection condition (which does not offer the best spatial discrimination).

We have demonstrated novelty through a behavioural study of gaze supported by a situated display. Technical novelty shows that a display like this can be made by loosely draping retro-reflective cloth over a chair. This ability to rapidly deploy a cost effective solution for normal spaces was specifically identified by BBC R&D. This could not have been achieved, for example, using a lenticular display which relies on retaining reflection of light paths so would have to be carefully installed set-up and aligned.

The null result for multi-view does not necessarily suggest that this element of the technology failed. It may suggest that task (in this case directional play in poker) can have more of an influence on gaze and mutual gaze than the supporting technology. While this might be intuitive we have not seen a similar study that provides this rigour of evidence. Particular attention was given to testing this null result as it is perhaps more interesting than finding the expected. While it seemed that there was no breakdown in normal conversation observable in the video playback, there is a difference in the questionnaire responses between face-to-face and technology mediated interaction. Participants felt that the Telethrone projection was 'in the space' with them in both projector conditions. Participants felt more like they were looking through a screen toward Telethrone in both conditions. Participants felt that the Telethrone user occupied the same space to them in both conditions. However participants more like they were looking into another room when looking at Telethrone in both conditions. Remote users seem naturally included, and support for multiple viewpoints onto a simple and affordable situated display has been technically demonstrated. Neither system seems to demonstrate obvious uncanniness at first inspection.

8. Future Work

We argue that further testing of the Telethrone is warranted adopting different tasks and/or experimental design. When examining the self-reported questionnaire responses there are significant differences between responses to the Telethrone when compared to the physically present researcher as show in Figure 11. A more differentiated control condition such as a standard webcam mediated session would better isolate this.

This research is an incremental step toward the kind of stereoscopic holographic effect popularly encountered in film and fiction. This is best exemplified in The High Council Chamber scene in Star Wars III:Revenge Of The Sith where some members are shown seated and tele-present but are viewable from any angle.

Research described in this paper provides a foundation for continuing research into projection on furniture. Multiple discrete viewpoints may show advantage over room2room, which is the other system currently demonstrating progress in projection onto furniture.

9. Acknowledgements

We wish to thank Claire Miller who assisted with the development of the Poker software, and Stephen Bowden who assisted with early work on Telethrone.

References

[AG76] ARGYLE M., GRAHAM J. A.: The central Europe experiment: Looking at persons and looking at objects. vol. 1, pp. 6–16. doi:10.1007/BF01115461. 2

[AI69] ARGYLE M., INGHAM R.: *Gaze, Mutual Gaze, and Proximity*. No. 1965. 1969. 2, 6

[AMEB12] AL MOUBAYED S., EDLUND J., BESKOW J.: Taming mona lisa: Communicating gaze faithfully in 2d and 3d facial projections. vol. 1, p. 25. 1

[Arg] ARGYLE M.: *Bodily communication (2nd ed.)*. 4

[BBBL01] BAILENSON J., BLASCOVICH J., BEALL A., LOOMIS J.: Equilibrium theory revisited: Mutual gaze and personal space in virtual environments. vol. 10, pp. 583–598. doi:10.1162/105474601753272844. 2

[BHPR] BOHANNON L. S., HERBERT A. M., PELZ J. B., RANTANEN E. M.: Eye contact and video-mediated communication: A review. doi:10.1016/j.displa.2012.10.009. 3

[BKKF13] BECK S., KUNERT A., KULIK A., FROEHLICH B.: Immersive group-to-group telepresence. vol. 19, IEEE, pp. 616–625. 2

[Can11] CANNY A. S. J.: Triple-view: Improving persuasion in group video conferencing through spatial faithfulness. 1, 2, 3, 7

[Ekm93] EKMAN P.: Facial expression and emotion. *American psychologist 48*, 4 (1993), 384. 2

[GBK05] GEE F., BROWNE W., KAWAMURA K.: Uncanny valley revisited. In *Robot and Human Interactive Communication, 2005. IEEE International Workshop on* (2005), Ieee, pp. 151–157. 3

[HH69] HALL E. T., HALL E. T.: *The Hidden Dimension*, vol. 1990. Anchor Books New York, 1969. 4, 5

[KBG*12] KIM K., BOLTON J., GIROUARD A., COOPERSTOCK J., VERTEGAAL R.: Telehuman: Effects of 3d perspective on gaze and pose estimation with a life-size cylindrical telepresence pod. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems* (2012), ACM, pp. 2531–2540. 3

[KCL13] KRISTOFFERSSON A., CORADESCHI S., LOUTFI A.: A review of mobile robotic telepresence. *Advances in Human-Computer Interaction 2013* (2013), 3. 3

[Kle86] KLEINKE C. L.: Gaze and eye contact: a research review. *Psychological bulletin 100*, 1 (1986), 78. 2

[LT11] LEE M. K., TAKAYAMA L.: Now, i have a body: Uses and social norms for mobile remote presence in the workplace. In *Proceedings of the 2011 annual conference on Human factors in computing systems* (2011), ACM, pp. 33–42. 3

[LWN*09] LINCOLN P., WELCH G., NASHEL A., ILIE A., STATE A., FUCHS H.: Animatronic shader lamps avatars. In *Mixed and Augmented Reality, 2009. ISMAR 2009. 8th IEEE International Symposium on* (2009), IEEE, pp. 27–33. 3

[Mor70] MORI M.: The uncanny valley. *Energy 7*, 4 (1970), 33–35. 3, 7

[MSSL*11] MITCHELL W. J., SZERSZEN SR K. A., LU A. S., SCHERMERHORN P. W., SCHEUTZ M., MACDORMAN K. F.: A mismatch in the human realism of face and voice produces an uncanny valley. *i-Perception 2*, 1 (2011), 10. 3

[Mun13] MUNSELL P.: Skype and messenger coming together: The next chapter. 2

[NC07] NGUYEN D., CANNY J.: Multiview: improving trust in group video conferencing through spatial faithfulness. pp. 1465–1474. 2, 3

[NC09] NGUYEN D. T., CANNY J.: More than face-to-face: empathy effects of video framing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2009), ACM, pp. 423–432. 2

[Ngu05] NGUYEN D.: Multiview : Spatially faithful group video conferencing. 2, 3

[OSS12] OYEKOYA O., STEPTOE W., STEED A.: Sphereavatar: A situated display to represent a remote collaborator. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2012), pp. 2551–2560. 3

[Ost14] OSTREM C.: A behavior analytic perspective on gambling behavior. Høgskolen i Oslo og Akershus. 5

[PS12] PAN Y., STEED A.: Preserving gaze direction in teleconferencing using a camera array and a spherical display. In *3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 2012* (2012), IEEE, pp. 1–4. 3

[PS14] PAN Y., STEED A.: A gaze-preserving situated multiview telepresence system. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems* (2014), ACM, pp. 2173–2176. 1, 3

[PSS14] PAN Y., STEPTOE W., STEED A.: Comparing flat and spherical displays in a trust scenario in avatar-mediated interaction. In *Proceedings of the annual ACM conference on Human factors in computing systems* (2014), ACM, pp. 1397–1406. 3

[RWLB01] RASKAR R., WELCH G., LOW K.-L., BANDYOPADHYAY D.: Shader lamps: Animating real objects with image-based illumination. In *Rendering Techniques 2001*. Springer, 2001, pp. 89–102. 3

[SKO*07] SAKAMOTO D., KANDA T., ONO T., ISHIGURO H., HAGITA N.: Android as a telecommunication medium with a human-like presence. In *Human-Robot Interaction (HRI), 2007 ACM/IEEE International Conference on* (2007), IEEE, pp. 193–200. 3

[SSA*01] SCHROEDER R., STEED A., AXELSSON A.-S., HELDAL I., ABELIN Å., WIDESTRÖM J., NILSSON A., SLATER M.: Collaborating in networked immersive spaces: as good as being there together? *Computers & Graphics 25*, 5 (2001), 781–788. 2

[STE*05] SCHREER O., TANGER R., EISERT P., KAUFF P., KASPAR B., ENGLERT R.: Real-time avatar animation steered by live body motion. In *Image Analysis and Processing–ICIAP 2005*. Springer, 2005, pp. 147–154. 3

[Tac03] TACHI S.: Telexistence and retro-reflective projection technology (rpt). In *Proceedings of the 5th Virtual Reality International Conference (VRIC2003) pp* (2003), vol. 69, pp. 1–69. 3

[TDYU11] TSUI K. M., DESAI M., YANCO H. A., UHLIK C.: Exploring use cases for telepresence robots. In *Human-Robot Interaction (HRI), 2011 ACM/IEEE International Conference on* (2011), IEEE, pp. 11–18. 3

[TP16] TOMISLAV PEJSA JULIAN KANTOR H. B.: Room2room: Enabling life-size telepresence in a projected augmented reality environment. ACM, pp. 1716–1725. 7

[Var02] VARIOUS: *Distributed Work*. MIT Press, 2002. 2, 3

[VD02] VERTEGAAL R., DING Y.: Explaining effects of eye gaze on mediated group conversations:: amount or synchronization? pp. 278–285. 2

[VVV00] VERTEGAAL R., VEER G. V. D., VONS H.: Effects of gaze on multiparty mediated communication. 5, 7

[VW03] VERTEGAAL R., WEEVERS I.: GAZE-2: conveying eye contact in group video conferencing using eye-controlled camera direction. *Proceedings of the SIGCHI . . .*, 5 (2003), 521–528. 2

[WRM*08] WOLFF R., ROBERTS D., MURGIA A., MURRAY N., RAE J., STEPTOE W., STEED A., SHARKEY P.: Communicating Eye Gaze across a Distance without Rooting Participants to the Spot. Ieee, pp. 111–118. doi:10.1109/DS-RT.2008.28. 2

[Zha07] ZHANG C.: Multiview imaging and 3DTV. *IEEE Signal Processing Magazine 1053*, 5888/07 (2007). 3